

КОМП'ЮТЕРНА СИСТЕМА РОЗПІЗНАВАННЯ ЖЕСТОВИХ КОМАНД КОРИСТУВАЧА З ВЕБІНТЕРФЕЙСОМ

Вінницький національний технічний університет

Анотація.

У представленій роботі розглянуто принципи побудови комп'ютерної системи розпізнавання жестових команд користувача з вебінтерфейсом на основі бібліотек комп'ютерного зору MediaPipe Hands та OpenCV. Описано архітектуру системи, алгоритм класифікації жестових команд на основі аналізу координат ключових точок кисті руки та механізм стабілізації результатів розпізнавання. Реалізовано вебінтерфейс із використанням HTML, CSS та JavaScript. Проведено тестування точності розпізнавання жестів та швидкодії системи.

Ключові слова: комп'ютерний зір, розпізнавання жестів, вебінтерфейс, MediaPipe, OpenCV, WebRTC, обробка відеопотоку, жестові команди, класифікація жестів.

Abstract.

This paper examines the principles of designing a computer system for recognizing user gesture commands via a web interface, based on the MediaPipe Hands and OpenCV computer vision libraries. The paper describes the system architecture, an algorithm for classifying hand gestures based on the analysis of key hand point coordinates, and a mechanism for stabilizing recognition results. A web interface was implemented using HTML, CSS, and JavaScript. Tests were conducted to evaluate gesture recognition accuracy and system performance.

Keywords: computer vision, gesture recognition, web interface, MediaPipe, OpenCV, WebRTC, video stream processing, gesture commands, gesture classification.

Вступ

Сучасні інформаційні технології активно впроваджуються у різні сфери діяльності людини, забезпечуючи зручнішу та ефективнішу взаємодію користувача з комп'ютерними системами. Одним із перспективних напрямів розвитку людино-машинної взаємодії є використання систем розпізнавання жестових команд, які дозволяють здійснювати безконтактне керування програмним забезпеченням і цифровими пристроями без традиційних засобів введення інформації.

Розвиток бібліотек комп'ютерного зору дозволив реалізовувати подібні системи на основі звичайних вебкамер без необхідності застосування спеціалізованих сенсорів.

Метою роботи є розроблення комп'ютерної системи розпізнавання жестових команд користувача з вебінтерфейсом та дослідження можливостей підвищення точності й швидкості розпізнавання жестів у режимі реального часу.

Основна частина

Розроблена система складається з модуля захоплення відеопотоку, модуля визначення положення руки, класифікатора жестових команд та вебінтерфейсу користувача. Отримання відеопотоку з вебкамери здійснюється засобами WebRTC, що забезпечує роботу системи безпосередньо у браузері без встановлення додаткового програмного забезпечення. Кожен кадр передається до модуля обробки зображення, де виконується попередня фільтрація та нормалізація відеоданих.

Для визначення положення кисті руки використано бібліотеку MediaPipe Hands, яка дозволяє визначати 21 ключову точку кисті та пальців у режимі реального часу. Отримані координати ключових точок передаються до модуля класифікації жестових команд.

Алгоритм класифікації жестових команд на основі аналізу координат ключових точок кисті руки виконується у декілька етапів, як наведено на рисунку 1.

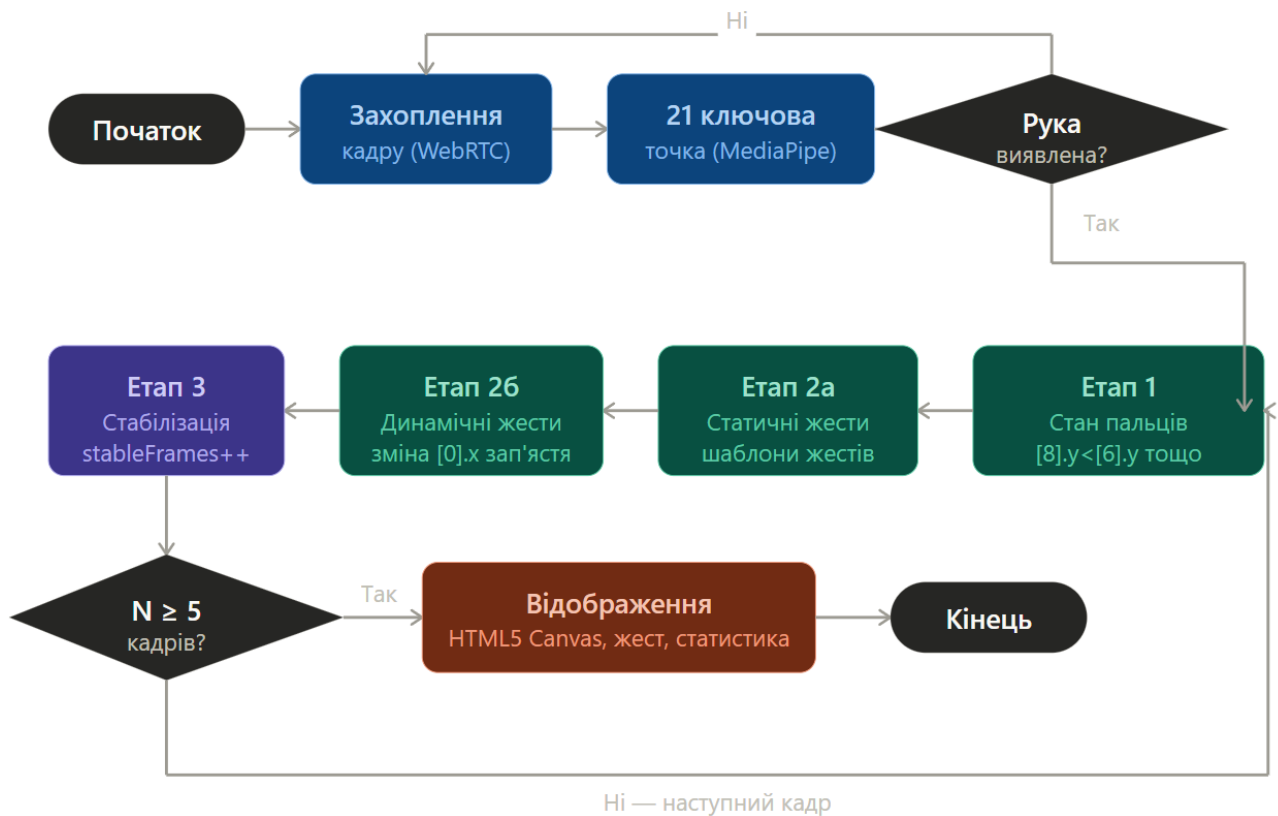


Рисунок 1 – Алгоритм класифікації жестових команд на основі аналізу координат ключових точок кисті руки

На першому етапі визначається стан кожного пальця: система порівнює координату кінчика пальця з координатою його середнього суглоба — якщо кінчик розташований вище за суглоб, палець вважається піднятим. Зокрема, для вказівного пальця перевіряється умова $landmarks[8].y < landmarks[6].y$, для середнього — $landmarks[12].y < landmarks[10].y$, для безіменного — $landmarks[16].y < landmarks[14].y$, для мізинця — $landmarks[20].y < landmarks[18].y$, а для великого пальця аналізується горизонтальне зміщення: $landmarks[4].x < landmarks[3].x$.

На другому етапі отримана комбінація стану пальців порівнюється із шаблонами жестових команд. Наприклад, усі підняті пальці відповідають жесту «Відкрита долоня», піднятий лише великий палець — жесту підтвердження, піднятий лише вказівний палець — жесту вибору елемента. Для розпізнавання динамічних жестів (рух руки вправо або вліво) додатково аналізується зміна координати зап'ястя між послідовними кадрами.

На третьому етапі виконується стабілізація результатів розпізнавання. Оскільки під час швидких рухів руки або часткової втрати руки із поля зору камери можуть виникати короточасні помилки класифікації, жестова команда підтверджується лише у випадку, якщо один і той самий жест стабільно визначається протягом не менше п'яти послідовних кадрів. Це дозволяє усунути випадкові помилкові спрацьовування та підвищити стабільність роботи системи.

Вебінтерфейс системи реалізовано із застосуванням HTML, CSS та JavaScript. Для відображення відеопотоку та накладання ключових точок руки використовується елемент HTML5 Canvas. Інтерфейс забезпечує відображення поточного розпізнаного жесту та статистики точності роботи системи.

Тестування системи проводилось для п'яти базових жестових команд по 50 повторень кожного в різних умовах освітлення та при різних відстанях між рукою та камерою. Встановлено, що оптимальною є відстань 50–70 см при стабільному освітленні та однотонному фоні зображення. Аналіз швидкодії показав, що система підтримує частоту обробки 24–30 кадрів за секунду при роздільній здатності відео 640×480 пікселів, а середній час реакції на жестову команду становить 80–120 мс.

Висновки

Розроблено комп'ютерну систему розпізнавання жестових команд користувача з вебінтерфейсом на основі бібліотеки MediaPipe Hands та технологій WebRTC і HTML5. Обґрунтовано вибір програмних засобів, що дозволив реалізувати систему без використання спеціалізованого апаратного забезпечення. Розроблено алгоритм класифікації жестових команд на основі аналізу координат 21 ключової точки кисті руки, який включає визначення стану кожного пальця шляхом порівняння координат кінчиків із координатами суглобів, порівняння отриманих комбінацій із шаблонами жестів та підтвердження команди після стабільного розпізнавання протягом п'яти послідовних кадрів. Перспективами подальшого розвитку є розширення набору підтримуваних жестів, використання нейронних мереж для підвищення точності класифікації та підтримка динамічних жестів.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Bradski G., Kaehler A. Learning OpenCV 4 Computer Vision with Python. — Birmingham : Packt Publishing, 2020. — 1024 p.
2. Szeliski R. Computer Vision: Algorithms and Applications. — Cham : Springer, 2022. — 979 p.
3. MediaPipe Hands Documentation. Google Developers. — Режим доступу: <https://developers.google.com/mediapipe>
4. OpenCV Documentation. — Режим доступу: <https://opencv.org/>
5. Mozilla Developer Network. WebRTC API. — Режим доступу: <https://developer.mozilla.org/>

Колесник Ірина Сергіївна — кандидат технічних наук, доцент кафедри обчислювальної техніки, Вінницький національний технічний університет, Вінниця, e-mail: iskolesnyk@gmail.com

Романенко Іван Володимирович — студент групи ІСП-22б, факультет інформаційних технологій та комп'ютерної інженерії, Вінницький національний технічний університет, Вінниця, e-mail: romanenkoivan20@gmail.com

Kolesnyk Iryna Serhiivna — Candidate of Technical Sciences, Associate Professor of the Department of Computer Engineering, Vinnytsia National Technical University, Vinnytsia, e-mail: iskolesnyk@gmail.com

Romanenko Ivan Volodymyrovych — student of group 1SP-22b, Faculty of Information Technologies and Computer Engineering, Vinnytsia National Technical University, Vinnytsia, e-mail: romanenkoivan20@gmail.com